UDC 534.86
621.395.9

1969
THE QUEEN'S AWARD
TO INDUSTRY

RESEARCH DEPARTMENT REPORT

Sound-quality improvement of broadcast telephone calls

No. 1972/26

Research Department, Engineering Division
THE BRITISH BROADCASTING CORPORATION

RESEARCH DEPARTMENT

SOUND-QUALITY IMPROVEMENT OF BROADCAST TELEPHONE CALLS

Research Department Report No. **1972/26**

UDC 534.86
621.395.9

It uses SI units in accordance with B.S. document PD 5686

M.G. Croll, B.Sc., A.R.C.S.

Head of Research Department

(EL-67)

Research Department Report No. 1972/26

SOUND-QUALITY IMPROVEMENT OF BROADCAST TELEPHONE CALLS

(EL-67)

August 1972

Research Department Report No. **1972/26**
UDC 534.86
621.395.9

SOUND-QUALITY IMPROVEMENT OF BROADCAST TELEPHONE CALLS

Summary

Several devices have been investigated which were suggested as means of improving the acoustical quality of telephone contributions in broadcast programmes. A promising processing system was simulated which included means of modifying the spectrum of telephone signals after receipt, a device to synthesise low frequencies, and some compensation for the absence of high frequencies. However, subjective tests showed that such a system effected only a small improvement in the quality of telephone signals. Although no complete processing system can therefore be recommended at this stage, some of the devices described in this report could form essential parts of a system in which other devices are available to reduce non-linear distortion and noise.

1. Introduction

1.1. General

Telephoned contributions from professional correspondents are frequently broadcast in news and current affairs programmes on both radio and television, and give a sense of actuality and immediacy to reporting. Telephone signals are also broadcast during programmes where listeners are invited to telephone the studio and contribute to a programme which is often 'live'. In both cases the telephone service is a valuable facility in that programme contributions can be accepted from any of the millions of telephones throughout the world. However, the acoustical quality of speech from the telephone network is poor by broadcasting standards, and the extensive use of telephoned contributions, particularly in some news programmes, has given rise to complaints from the general public.

Most telephone calls that are broadcast at present are intelligible, but understanding them requires more effort by the listener than does high-quality studio speech. The characteristic sound of telephone speech is unnatural and, when interposed between passages of studio speech, it can become distracting because understanding the two completely different types of speech requires different levels of concentration. Therefore, in processing telephone speech to improve its quality, efforts should be made to make the processed speech quality closer to that of the studio speech. Moreover, because certain foreign correspondents become well known through the contributions they frequently make, it is desirable that any processing should not destroy the identity of the original voice. Ideally, processed speech should sound the same as studio speech from a given person.

This report considers a number of processes that could be applied to speech signals from the telephone network to improve their suitability for broadcasting. Obtaining the basic electrical signals from telephone circuits and controlling them in level is considered in the Appendix to this report. Ways and means of improving telephone speech from frequent broadcasters at the sending end are not considered here; BBC Radio Broadcasting Operations and Maintenance staff and Designs Department are working on this problem.

1.2. The impairments and distortions of a telephone system

The telephone handset microphone is frequently the main distorting element in the telephone system. The main distortions it introduces are a non-uniform and band-limited amplitude/frequency response, a non-linear phase/frequency characteristic and various forms of input/output non-linearity. The electrical signals generated are then subjected to further amplitude/frequency and phase/frequency distortions in the telephone network circuits where noise is also added to the signal. The bandwidth of the signal is usually limited to 300 Hz to 3·4 kHz, although some circuits have an upper frequency limit of 2·5 kHz. In this investigation attempts were made to reduce the effects of amplitude/frequency distortion, restricted bandwidth and noise. Although non-linear distortions cause a major impairment of the signal, means of reducing them in the presence of other distortions remain to be devised in possible future work.

2. Correction of amplitude/frequency distortion*

2.1. Equalisation requirements

The cumulative effects of the amplitude/frequency responses of different parts of the telephone system impair speech quality and may often effectively reduce the passband to considerably less than the nominal 300 Hz to 3·4 kHz. For the present purpose it is therefore desirable to

* Certain proposals in this section are due to R.N. Robinson and a patent has been applied for.[1]

equalise the amplitude/frequency characteristic of telephone signals within the nominal passband. As well as improving the audio signal quality, such equalisation facilitates other forms of processing which will be described later.

Equalisation of the response of systems is usually carried out with the aid of test signals which are applied to the system input and measured at the output, but in the case being treated here, the only signals applied to the system input are the speech signals generated by the caller. However, within the telephone bandwidth the speech-signal spectrum measured in 1/3-octave bands is similar for all people, both male and female. When analysed in this way, the spectrum of studio-quality speech signals usually lies within ±3 dB of the average speech spectrum, whereas the spectrum of telephone-speech signals deviates by up to 20 dB from the average speech spectrum. Therefore, the matching of the output spectrum with the average speech spectrum was chosen as a basic requirement for equalisation of telephone-speech signals for broadcasting.

2.2. Practical methods of equalisation

Equalisation of the telephone signal by filtering it into 10 1/3-octave bands, measuring the signal levels, and independently adjusting the gain in each band would be a time-consuming and costly process. However, after analysing the spectra of a number of telephone signals, it was found that the required equalisation could usually be achieved using the two fixed-frequency variable-depth notch-filter characteristics shown in Fig. 1.

The way in which these notch-filters were used to equalise a telephone signal is illustrated in Fig. 2. Using only two variable parameters, namely the notch-filter depths, it was found that the spectra of telephone calls could usually be matched to within ±3 dB of the average speech spectrum.

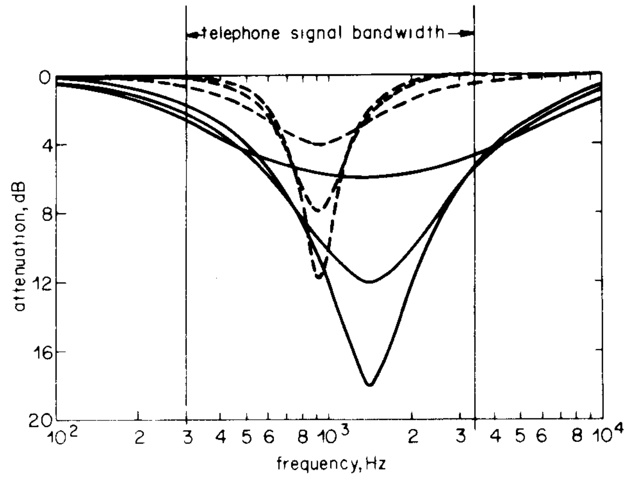

Fig. 1 - Amplitude/frequency characteristics of the notches

(*a*) Notch A	Centre frequency 920 Hz
– – – –	Depths 4, 8, 12 dB shown
(*b*) Notch B	Centre frequency 1·4 kHz
———	Depths 6, 12, 18 dB shown

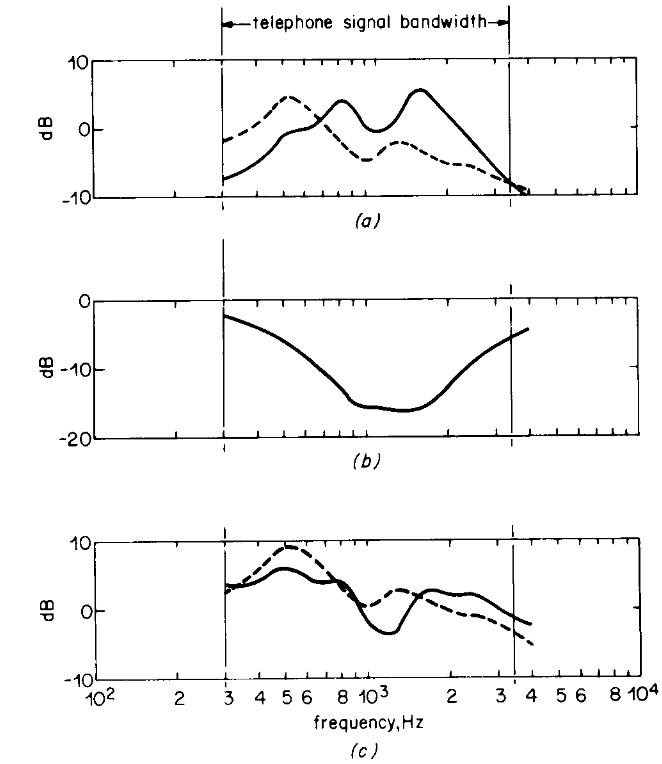

Fig. 2 - Equalisation of telephone speech signals

(*a*) ——— Telephone speech spectrum
– – – – – Average voice spectrum

(*b*) Equalisation characteristic:
4 dB deep 920 Hz Notch A
14 dB deep 1·4 kHz Notch B

(*c*) ——— Equalised telephone speech spectrum
– – – – – Average voice spectrum

2.3. Adaptive operation of equaliser

A device was built to adjust the depth of one fixed-frequency notch to prove the feasibility of an adaptive equaliser meeting the requirement stated in Section 2.1. Although the device was capable of only partial equalisation, it served to indicate whether any unsurmountable fundamental or technological problems would be encountered with an automatic full-equaliser.

A schematic of the system is shown in Fig. 3. The depth of the notch characteristic is controlled by an error signal generated in a feedback control loop. The output from the equaliser is first applied to a spectrum-shaping circuit whose amplitude/frequency response is the inverse in amplitude of the average speech spectrum. The output from this circuit is then applied to the inputs of complementary bandpass and bandstop filters, both with centre frequencies equal to the centre frequency of the equaliser notch and bandwidths of 900 Hz. Measuring circuits connected to the outputs of the bandpass and bandstop filters determine the average amplitudes of the signal in the region of the notch centre frequency and in the remainder of the telephone bandwidth. These average amplitudes are then compared and their difference gives rise to an error signal for the voltage-controlled equaliser. Hence the action of the equaliser is to maintain the spectrum of the output

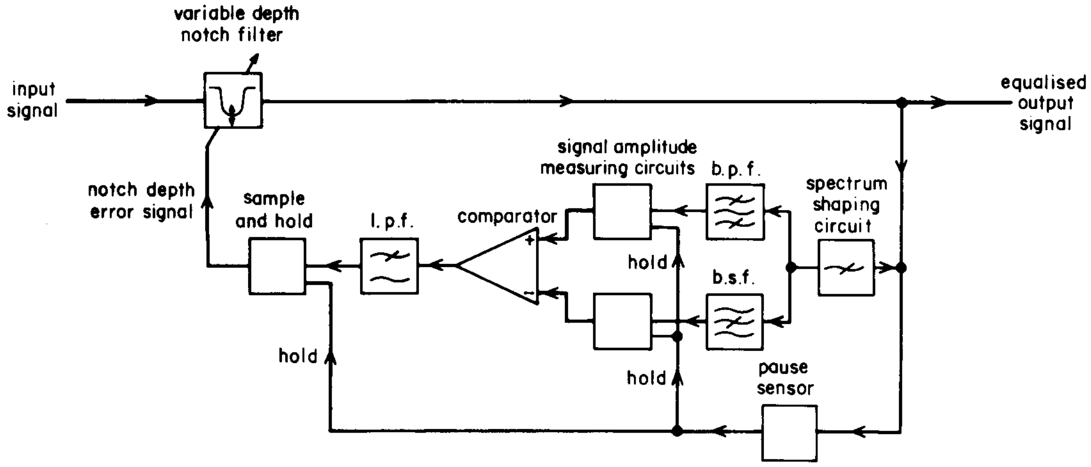

Fig. 3 - Schematic of automatic equaliser

signal similar to the average speech spectrum by attempting to make a flat amplitude/frequency spectrum at the output of the spectrum-shaping circuit. To ensure that the system operates only on speech components a pause sensor (described more fully in the Appendix, Section 9.4.3) controls circuits which freeze the amplitude measurements and the depth of the notch when the envelope of the speech signal falls below a prescribed level.

Time constants in the feedback control loop were found to be critical. If the equaliser operated too fast the amount of equalisation varied significantly from word to word for one caller producing subjectively distracting effects. However, it was desirable for the equaliser to operate as quickly as possible when a telephone call was first connected. With the time constants set to give only a ±3 dB fluctuation of notch depth during any one call, the equaliser took about 15 seconds to slew itself from one end of its range to the other when an appropriately different telephone signal was applied. When the device was applied to the equalisation of a series of telephone speech and studio-quality speech excerpts the only subjectively undesirable effects were due to the rather long slewing time of 15s. It is possible that the slewing time could be reduced by using more sophisticated circuits but it is unlikely that times of less than 5s could be achieved while maintaining the ±3 dB fluctuation referred to above. If the long slewing time proved to be a problem in practice, the equalisation could be established during some preliminary speech prior to the telephone call being contributed to the programme. The amount of equalisation could then be maintained constant for the duration of the call.

3. Extension of the bandwidth

3.1. General

The telephone bandwidth has a lower limit of about 300 Hz and an upper limit of about 3·4 kHz. Although it is possible that some additional speech components from outside this bandwidth could be retrieved by equalisation, there would be an overall signal impairment due to the consequent increase in noise. Therefore, if telephone speech signals are to be processed to become similar to high-quality speech, the processing must include some means of synthesising components outside the telephone bandwidth.

The final solution to the problem of increasing the bandwidth of the signal probably lies in the use of an analyser and synthesiser to break down the input signal into its constituent speech parameters, modify the parametric values, and recombine the parameters as wideband speech-signals. There are research workers who have reached advanced stages in developing real-time speech analysers and synthesisers[2] and who might eventually be able to provide the basic equipment needed to facilitate bandwidth extension and other processing of telephone-speech signals. However a full investigation of such techniques is outside the scope of the present investigation. Here we have attempted to develop techniques, based on simple forms of analysis, which could be used to generate components outside the telephone bandwidth for addition to the telephone-speech signal.

3.2. Lower-frequency synthesis

The speech sounds made by the mouth can be classified as either voiced or unvoiced sounds. The unvoiced sounds are the plosives and sibilants, and they account for most of the high-frequency energy of speech; these will be considered later. The voiced sounds are the vowels and the main part of some consonants, and they account for almost all of the speech energy at low frequencies. The voiced sounds are produced by the larynx and filtered by the action of the vocal tracts and the nasal cavity. The sound produced by the larynx has a pitch (or fundamental frequency) and is rich in harmonics. The vocal tracts act on this signal as a series of resonant circuits whose frequencies are independently variable. Each resonant mode of the vocal tracts is known as a formant. The nasal cavity has a fixed configuration, which produces some spectral shaping, and provides an additional outlet path for the sounds produced by the larynx. This path is opened and closed during the process of speech production.

We can estimate the nature of the original speech components below 300 Hz from the telephone signal components and a knowledge of the normal amplitude spectra of speech parameters.[3,4]

The fundamental frequency of the larynx is in the range 80 Hz to 200 Hz for most speakers. The lowest formant frequency is usually within the range 200 Hz to 1 kHz while other formant frequencies are usually above 800 Hz. Hence there could be the fundamental frequency and up to three harmonics falling below the telephone bandwidth. The relative amounts of energy at these frequencies depends on the activity of the larynx and is influenced by the effect of the nasal cavity and, less frequently, by the lowest formant.

In this investigation devices were considered where the fundamental frequency and its harmonics were synthesised in fixed proportions, and there was no adaptation to take account of variations due to the nasal and first-formant effects. The most successful device was based on the well established principle that the speech envelope contains a large component at the fundamental frequency.[5] It was found that when telephone-speech signals were full-wave rectified and band-pass filtered from 80 Hz to 300 Hz, the output contained components at the fundamental frequency and its harmonics. Attempts to extract solely the fundamental frequency, i.e. the pitch, were more successful with telephone-band-restricted high-quality speech than they were with telephone-speech signals.

The success of the rectifying and filtering technique in synthesising realistic low-frequency speech components depended critically on the quality of the prevailing input signal, and the results were improved when telephone signals were equalised before being applied to the synthesiser. Some of the difficulty experienced in extracting the fundamental frequency from telephone signals was due to the phase/frequency and non-linear distortions present in the input signal.

However, using these techniques, it was found possible to generate some low frequencies which could be added with advantage to equalised telephone-speech signals for broadcasting. This scheme is remarkably similar to one proposed in Germany in 1933 for enhancing the acoustical quality of telephone speech at frequencies below 600 Hz.[6]

3.3. Higher-frequency synthesis

The main components of speech energy at frequencies above the telephone bandwidth are unvoiced and occur during sibilants and plosives. These noise-like sounds are produced by the passage of air through a constriction formed between the tongue and either the roof of the mouth or the teeth. The spectrum of the noise is modified by the action of the vocal tracts to produce the different sibilant sounds.

As stated in Section 3.2, the main speech energy components within the telephone band are voiced, but there is also a significant contribution from these components at higher frequencies. Therefore, because of the different proportions and natures of the voiced and unvoiced components, it is necessary to differentiate between them in order to generate realistic higher-frequency speech components.

Extreme difficulty was experienced in constructing a simple voiced/unvoiced component detector which would work reliably for high-quality telephone-band-restricted speech. The most successful of the devices tried was one which compared the speech energy in the upper part of the telephone band with that in the lower part. This detected most sibilants but incorrectly classified the 'e' sound as unvoiced. When applied to equalised telephone signals, this device was even less successful and tended to operate randomly, probably because non-linear distortions caused significant and spurious changes in the spectrum of the telephone speech.

Hence, in this investigation no simple means of generating high frequencies was discovered which could be used with advantage. Additional components generated from both voiced and unvoiced signals without differentiating between them caused further impairment when added to the original telephone signal. Those generated using an unreliable voiced/unvoiced component detector were distracting and reduced the intelligibility of the signal.

The best that could be done was to compensate for the lack of high frequencies by providing a boost to components in the upper part of the telephone bandwidth. The amplitude/frequency characteristic of the boost is shown in Fig. 4. It was found that for most telephone calls a 10 dB boost could be applied with advantage.

4. Noise reduction systems

The unwanted noise generated within a telephone system contains both impulsive and random components. For telephone circuits which have a high attenuation, the final signal-to-noise ratio is quite unacceptable for broadcasting. Also, the action of the equalisers increases the noise level.

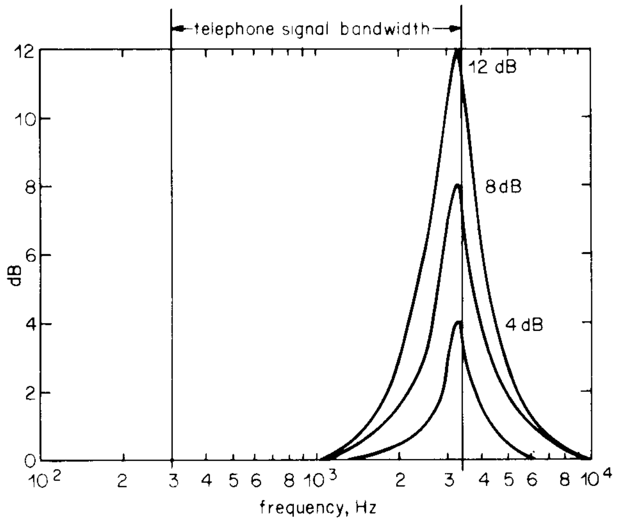

Fig. 4 - The variable high-frequency boost circuit

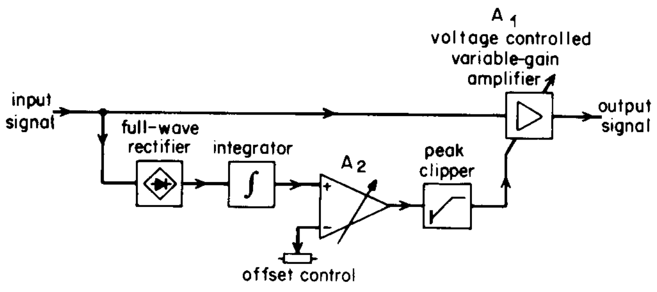

Fig. 5 - Schematic of expander/noise gate

Syllabic expanders and noise gates were investigated as means of reducing the noise from a telephone system. The general schematic for these devices is shown in Fig. 5. The gain of the variable-gain amplifier A_1 is determined by a signal derived from the syllabic components of the envelope of the input signal. If the gain of the side-chain A_2 is large the device operates as a noise gate. With lower gains of the amplifier A_2, the device operates as a syllabic expander. Typical transfer characteristics for the device are shown in Fig. 6.

It was found that unless the signal-to-noise ratio was already better than about 40 dB this device produced distracting effects. With some telephone calls, modulation of the noise by the signal was distracting; in other cases the incoming noise operated the expander and was thus modulated. If the threshold for the expander was set at too high an input level, the intelligibility of the speech was impaired.

Hence, no expander configuration was devised that could be used to reduce the impairment of telephone signals caused by noise, without introducing other impairments that were just as distracting. A solution to the problem of extracting speech signals from noise has been suggested[7] which involves the use of a complex signal analyser capable of distinguishing between speech and non-speech signals, and of cancelling the noise by means of a correlation technique.

5. Simulation of an automatic processing system

To test the effects of equalisation, lower-frequency synthesis and a boost of the high frequencies in the telephone-frequency band, these processes were applied to 19 telephone calls from the programme 'It's Your Line'. For each call the signal was first analysed into $^1/3$-octave bands using a spectrum analyser which displayed on an oscilloscope the logarithm of the peak signal level in each band. A time exposure photograph was made of the display to record the signal spectrum integrated over the duration of the telephone call. The amount of equalisation necessary for each call was then calculated and applied to the signal using the two variable-depth fixed-frequency notches described in Section 2.2 of this report. Synthetic low frequencies were then generated as described in Section

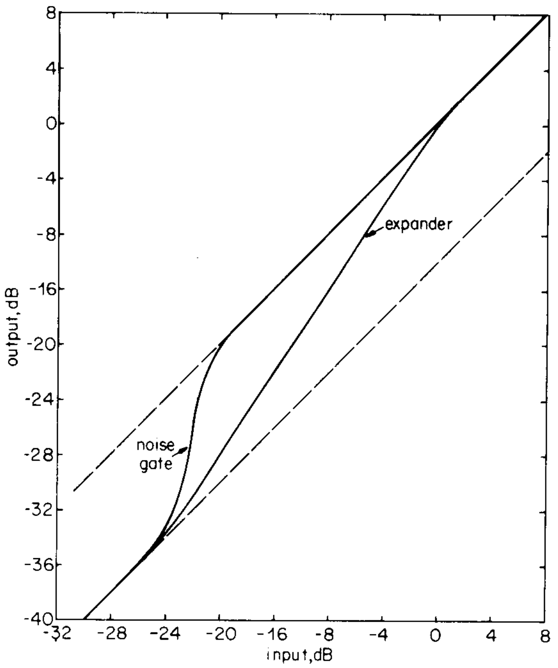

Fig. 6 - Transfer characteristic of expander and noise gate giving a 10 dB reduction of noise

3.2 and added to the signal, and the high frequencies were boosted as described in Section 3.3. The amounts of synthesised low-frequency signals and high-frequency boost added to the signal were adjusted and optimised by a small number of people who individually listened to the output using a high-quality loudspeaker. The listening room had a volume of $85m^3$ and a mid-band reverberation time of 0·3s. Twin-track recordings were made of each telephone call with the processed signal on one track and the original (unprocessed) signal on the other.

The results after processing the telephone signal were thought to be an improvement although the quality of the processed signal was still poor by comparison with high-quality studio speech. To obtain a quantitative assessment of the processed signal quality, a series of subjective tests were conducted, as described below.

6. Subjective tests

6.1. General

Three series of tests were carried out to evaluate the effect that processing the telephone-speech has on its intelligibility, ease of understanding, and annoyance. Material for these tests was selected from the recordings made of the processed and unprocessed telephone calls to 'It's Your Line'. The tests were conducted in the listening room using a high-quality loudspeaker with groups of 6 observers at a time. A total of twenty-four people participated in each series of tests; twelve were from the scientific staff of Research Department and were experienced in

assessing speech signals, and the other twelve were members of the non-scientific staff. In each series of tests the material was presented both processed and unprocessed, and each individual test contained almost equal numbers of processed and unprocessed items in random order.

6.2. Word intelligibility

The word intelligibility of the signals was measured by presenting 88 single words in random order and asking each observer to write down the words he understood. For these tests only words which were difficult to understand were selected and care was taken to ensure that observers were unlikely to memorise the list of words.

The total scores for this series of tests were:

Unprocessed: 557 correct answers out of 2112,

Processed: 623 correct answers out of 2112.

This shows a very slight average improvement in the word intelligibility due to processing. No particular telephone call was significantly changed in word intelligibility when processed, and the average scores for scientific and non-scientific observers were similar.

6.3. Ease of understanding phrases

In this series of tests, 38 phrases were presented for assessment by the observers. Two short phrases were taken from each telephone item and presented in random order, both processed and unprocessed. The observers were asked to grade each phrase using the scale shown below, which is similar to one used by the Post Office for similar tests.

1. Complete relaxation possible — no effort required
2. Attention necessary — no appreciable effort required
3. Moderate effort required
4. Considerable effort required
5. Extreme effort required
6. Unintelligible

The results showed that the average grades for 'ease of understanding' were improved from 2·9 to 2·6 when the phrases were processed. Although small, this 0·3 of a grade improvement was probably significant as the standard deviation was 0·25 of a grade. The results from both the scientific and non-scientific observers showed good agreement, and, as the histogram in Figure 7 shows, few telephone calls were degraded when processed and 35% were improved by more than 0·5 of a grade.

6.4. Annoyance

This series of tests was intended to measure the annoyance of a telephone call when it was included in a programme which comprised mainly high-quality studio speech. Excerpts of telephone speech lasting about 10 secs were used, preceded by a similar duration of high quality studio speech the meaning of which told the observers

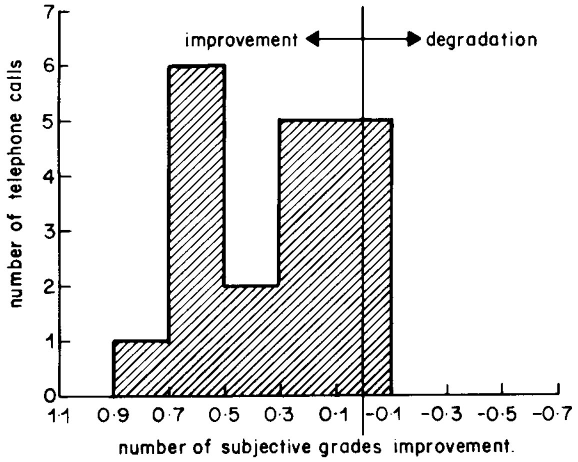

Fig. 7 - Effect on ease of understanding phrases

where on the score sheet they should write their assessment of the telephone call. The grades used in this series of tests were.

1. Not annoying
2. Slightly annoying
3. Moderately annoying
4. Definitely annoying
5. Very annoying
6. Intolerably annoying

The results showed that, on the average, processing the telephone signals increased their annoyance grade by 0·1 from 3·4 to 3·5 on the scale shown above. This increase is insignificant as the standard deviation was 0·3 of a grade. There was a very wide spread in the results, and the mean score of the scientific observers was 0·25 of a grade more favourable to the processed signal than that of the non-scientific observers. A histogram showing the mean results for different calls is shown in Fig. 8.

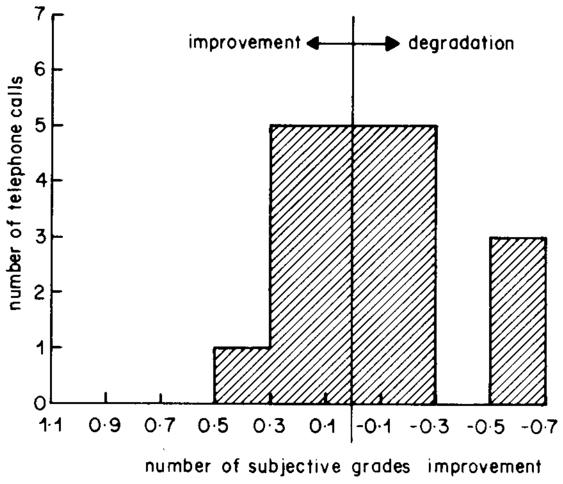

Fig. 8 - Effect of annoyance of telephone calls

6.5. Discussion of results

The results of the subjective tests showed that whereas processing should slightly improve the word intelligibility and ease of understanding phrases, their annoyance is not likely to be significantly reduced. This is difficult to explain, but comments made by some of the observers after the tests indicated that they were used to listening to broadcast telephone calls and that they sometimes found the unprocessed signals less annoying because they were able to adapt their senses to listening to them very quickly. Also, they likened the processed signals to very poor quality studio signals. Some observers found that the slight increase in noise level which occurred when the signal was equalised increased the annoyance.

For some of the processed items the synthetic low-frequency signals were not considered by the author to be very natural; for others, the equalisation and high-frequency boost increased the level of distortion products present in the original telephone signal. These effects might have accounted for some increase in the annoyance of a call, but their occurrence did not correspond with those calls which observers thought were more annoying when processed.

7. Conclusions and recommendations

In this work several devices were investigated which were suggested as means of improving the acoustical quality of telephone contributions in programmes. However, subjective tests showed that a system based on the most successful combination of these devices effected only a slight improvement in the acoustical quality of the telephone calls used in the tests.

Many telephone contributions which have been broadcast were apparently of poorer acoustical quality than those selected for the tests. These poorer-quality calls had high noise levels and a high percentage of non-linear distortion; neither of these two impairments were treated successfully in this investigation.

Experiments showed that appropriate matching of the telephone-speech spectrum to that of average speech can be adequately achieved with a pair of notch filters, and that a corresponding automatic spectrum equaliser is feasible. The latter could form an essential part of a more complex processing system.

The difficulty of generating acceptable wideband signals from telephone-bandwidth input signals became apparent during the work. If equalised wideband speech signals are required, the most promising technique for achieving this would appear to involve a combination of a formant-tracking speech analyser, a parametric automatic equaliser, and a parametric speech synthesiser, all operating in 'real-time'. For further developments on this front, the progress in research establishments at present investigating speech analysis and synthesis techniques should continue to be carefully monitored. Work in Research Department should be resumed if and when it becomes apparent that the results of the work in these establishments could be applied or adapted to improving telephone-speech quality for broadcasting.

8. References

1. British Patent Application No. 697/71 – Quality improvements in received telephone speech.

2. OLIVE, J.P. 1971. Automatic formant tracking by a Newton-Raphson technique. *J. acoust. Soc. Am.*, 1971, **50**, 2, (part 2), pp. 661 – 670.

3. FANT, C.G.M. 1960. Acoustic theory of speech production. 's Gravenhage, Moulton and Co., 1960.

4. FLANAGAN, J.L. 1965. Speech analysis synthesis and perception. Berlin, Springer Verlag, 1965.

5. SCHROEDER, M.R. 1970. Parameter estimation in speech: a lesson in unorthodoxy. *Proc. IEEE*, 1970, **58**, 5, pp. 707 – 712.

6. SCHMIDT, KARL-OTT-. 1933. Neubildung von underdückten Sprachfrequenzen durch ein nichtlinear verzerrendes Glied. *Telegraphen-und Fernsprech-Technik*, 1933, 1, p. 13.

7. SCHROEDER, M.R. 1971. Computers in acoustics. Proc. 7th International Congress on Acoustics, 1971, Vol. 1, p. 257.

8. ROSENBERGER, J.R. *and* THOMAS, E.T. 1971. Performance of an adaptive echo canceller operating in a noisy, linear, time-invariant environment. *Bell. Syst. tech. J.*, 1971, **50**, 3, pp. 785 – 813.

9. MITCHELL, O.M.M. *and* BERKLEY, D.A. 1971. A full-duplex echo suppresser using center-clipping. *Bell Syst. tech. J.*, 1971, **50**, 5, pp. 1619 – 1630.

10. SHORTER, D.E.L. *and* MANSON, W.I. 1969. The automatic control of sound-signal level in broadcasting studios. BBC Monogr., 1969, No. 77.

9. Appendix:
Apparatus to facilitate tests on improvements in broadcast telephone calls

9.1. Introduction

The apparatus described in this Appendix was built to facilitate tests on improvements in the sound quality of broadcast telephone calls, because of difficulty with the apparatus currently used at studios for direct connection to the telephone network.*

The apparatus separates the incoming telephone signal from other signals present on the telephone line and automatically controls the level to combat variations of the incoming speech-signal level over a 40 dB range.

9.2. Design of apparatus

The apparatus was designed to be connected to the two wires that normally feed a complete telephone set. These wires carry all the necessary signalling and operating voltages together with the incoming and outgoing speech signals. For our application a normal telephone set is used to establish contact with the remote caller. Then the apparatus is connected to the wires in place of the telephone set, to enable the person in the studio to converse with the caller.

Fig. 9 shows a block schematic of the apparatus. Signals from the studio microphone are amplified, spectrally shaped, and compressed to make them similar to telephone signals. They are then applied to the telephone circuits via a hybrid coil and an isolating transformer which is used to block the signalling and d.c. voltages present on telephone wires. Incoming telephone speech signals also pass through the isolating transformer and are applied to the hybrid coil. The incoming signal from the hybrid coil contains some proportion of the outgoing signal depending on the effectiveness of the coil. This separated incoming signal is band-pass filtered and applied to the automatic level controller. The links shown in the figure enable signal processing units which might affect the signal level to be inserted in such a way that their output level is controlled. Units which are sensitive to signal level can be inserted within the loop of the controller whilst those which are not can be inserted before the controller.

* More recently, improved studio equipment has been installed, which overcomes some of the difficulties experienced when using the old equipment.

The arrangements for separating the incoming from the outgoing signal and for controlling its level are described in greater detail in the following sections.

9.3. Separation of incoming from outgoing signals

9.3.1. General

Outgoing signals are normally applied to telephone circuits at a maximum level of about −5 dBm which has been found to be about 15 dB above the average level of incoming signals. Although the incoming signal level can vary by up to 20 dB from this average, it is seldom higher than that of the outgoing signal and can be up to 35 dB lower in level. If no measures are used to provide attenuation of the outgoing signal in the incoming signal path, problems of instability can arise and in some cases electrical and acoustic noises from the studio microphone can reach levels similar to that of the incoming telephone signal.

To completely suppress all outgoing signals in the incoming signal path would simplify automatic control of signal level. Some sophisticated devices have been designed

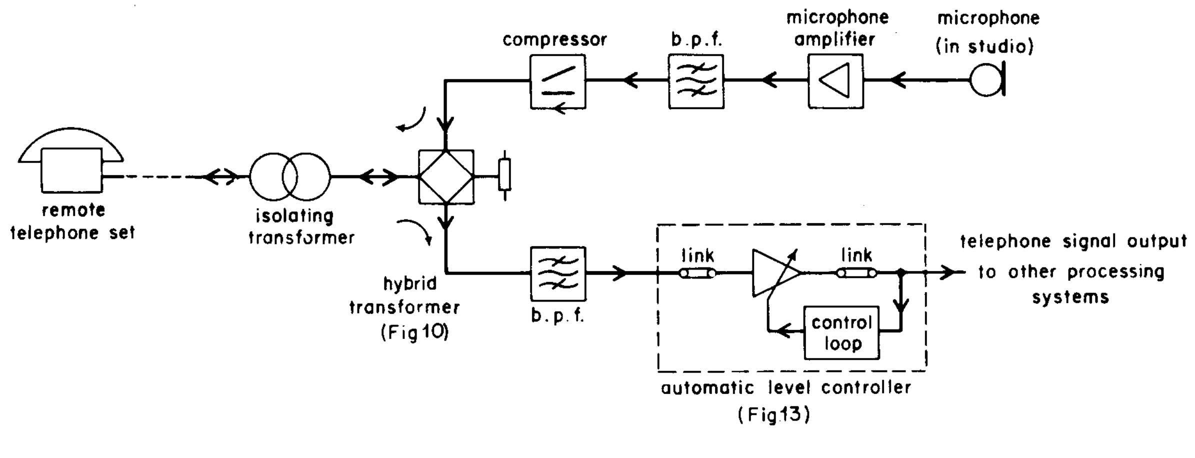

Fig. 9 - Schematic of apparatus

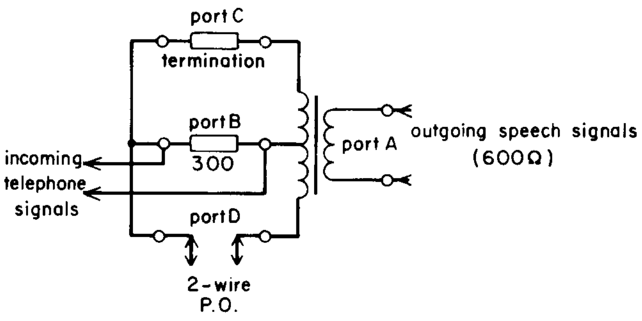

Fig. 10 - A 600 Ω hybrid coil

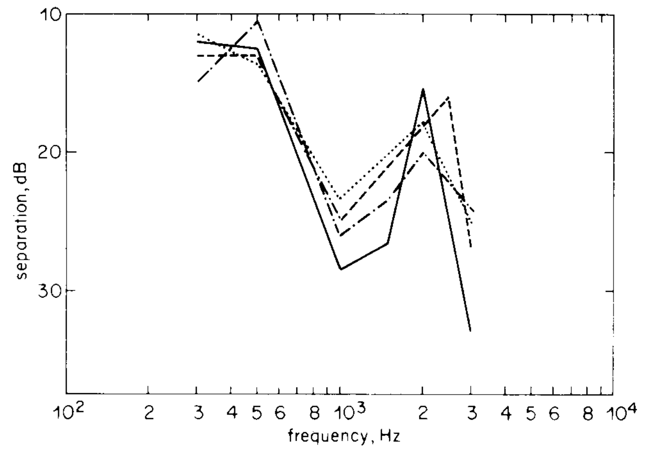

Fig. 11 - Measured separations achieved for four telephone lines from B.H. London

to do this in cases where long distance telephone lines cause echoes.[8,9] These devices are costly, however, and a full investigation of their performance for this application was beyond the scope of the present work which was limited to finding bow more conventional means (using a hybrid coil) could best be applied.

9.3.2. The hybrid coil

Fig. 10 shows a practical configuration of a 600 Ω hybrid coil, where no power is exchanged between ports A and B if the impedance at port C is equal to that at port D. If the impedances at C and D are not equal then some of the signal applied to A appears at B. From theory and confirmatory measurements, if a 20 dB separation is to be achieved, the resistive part of the terminations should be equal within about 20% and the phases of the impedances should be equal within about 10°, assuming that the terminations are approximately 600 Ω resistive.

The characteristic impedance of P.O. circuits is nominally 600 Ω. However, the cables are lossy and present a complex impedance; also series capacitors are used at exchanges to bridge between lines. Therefore in designing the termination impedance for a hybrid coil the reactive components of the impedance of a P.O. two-wire circuit must be taken into account.

With the help of the Post Office, the circuits from Broadcasting House, London, normally used for telephone contributions to discussion programmes, were traced, and networks simulating their impedances were used to terminate port C of the hybrid coil. These were tested in a practical set-up and a curve of the separations achieved for four telephone lines is shown in Fig. 11. The results show the separation achieved to be better at higher frequencies than at lower frequencies.

9.3.3. Improving the separation at lower frequencies

It was found that the separation at low frequencies could be improved, in effect, by applying a spectrum weighting network to the outgoing signal generated by the studio microphone. The amplitude/frequency response of a suitable network is shown in Fig. 12 and is similar to the average response of a telephone handset microphone. The quality of the studio signal transmitted to the distant caller is degraded when this network is used but tests have shown that the degradation is not significant and the intelligibility of the signal is not appreciably affected. By this means, separations of the order of 20 dB were achieved within the telephone band for most circuits.

9.4. Automatic level control

9.4.1. General

The signal level from a telephone system can vary by up to 40 dB depending on the caller and the attenuation of the P.O. circuits being used. Any automatic system of level control must be capable of handling these widely-varying signal levels without introducing any distracting effects. Moreover, the telephone speech signal is compressed by the action of the telephone microphone, and it

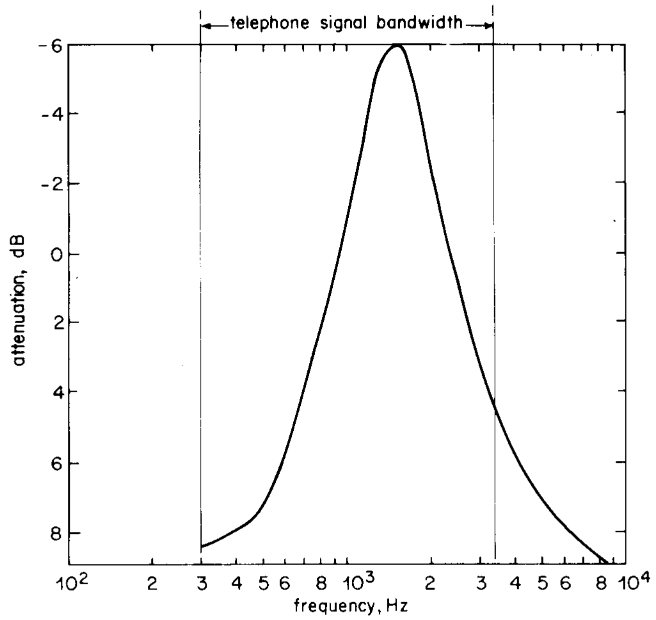

Fig. 12 - Amplitude/frequency characteristic of the spectrum shaping circuit

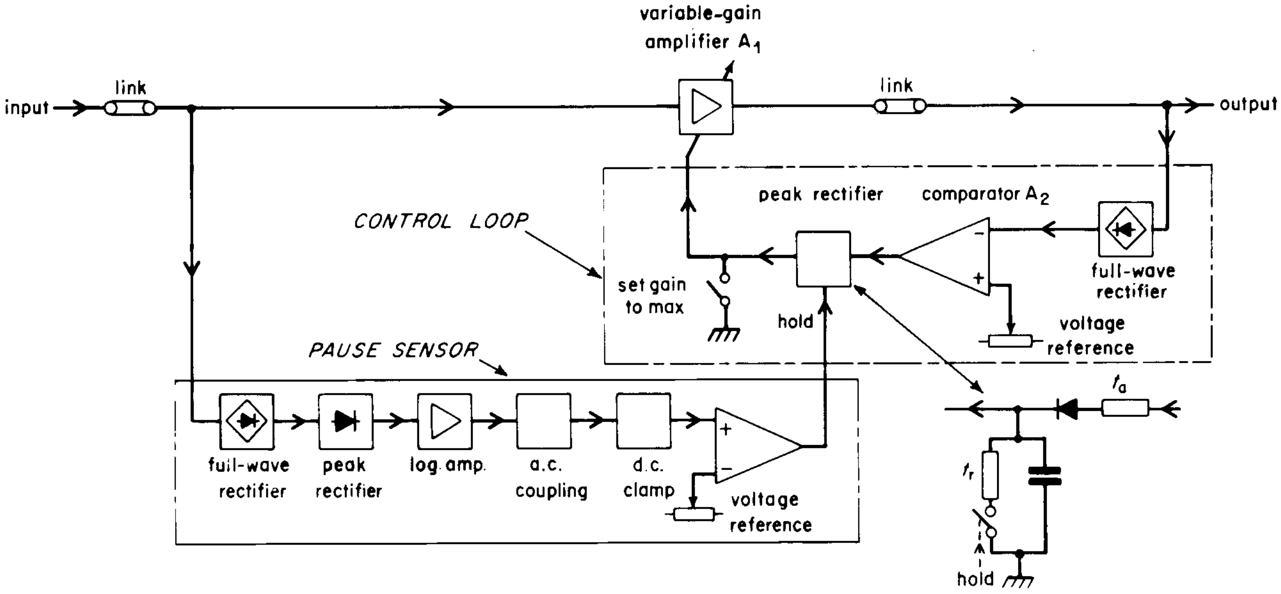

Fig. 13 - Schematic of automatic level controller

is desirable that the level controller should not introduce any further compression of the dynamic range of the telephone signal.

Limiters have been used for automatic level control at unattended studios,[10] but when they are used to control signals over a 40 dB range severe compression is introduced; also, during pauses in the signal, the gain is increased in such a way that noise is then amplified. To meet the requirements for automatic level control of telephone signals a more sophisticated device was developed. It includes circuits to hold the gain fixed during programme pauses.

9.4.2. Description of the controller

A schematic of the device is shown in Fig. 13. The input signals are applied to a voltage-controlled variable-gain amplifier, A_1. The output from this amplifier is the level-controlled output signal. Control signals which vary the gain of A_1 are generated in the control loop. Here the amplitude of the output signal is measured using a full-wave rectifier and a comparator A_2. The signal generated by A_2 is applied to a peak rectifier and the resultant signal is then used to vary the gain of A_1.

In the control loop, if the signal exceeds the reference voltage in A_2, the gain of A_1 is reduced with a time constant t_a determined in the peak-rectifier. When the signal does not exceed the reference voltage then the gain of A_1 is steadily increased with a time constant t_r, also determined in the peak-rectifier. This mode of operation is similar to that of a conventional limiter. However, in this device, the time constant t_r has two values t_{r1} and t_{r2} determined by a pause sensor. The shorter time constant t_{r1} is used when programme is present and t_{r2}, which is very long, is used during pauses.

9.4.3. The pause sensor

In the pause sensor the reference voltage corresponding to the pause-sensing threshold is automatically set, to a level which is a fixed number of decibels above the steady noise level of the signal.

A schematic of the pause sensor is included in Fig. 13. The signal envelope is measured using a full-wave rectifier and peak rectifier. It is then logarithmically amplified and suitably coupled to a d.c. clamp so that the part of the resultant waveform which corresponds to the steady noise level of the signal, is clamped to an arbitrary voltage. The signal is then compared with a reference voltage which corresponds to an input signal level about 10 dB above the noise level. A pause is signified when the signal is less than the reference voltage.

9.4.4. Performance of the automatic level controller

The performance of the automatic level controller was optimised by adjusting its time constant using recorded telephone-speech excerpts. The final time constants chosen were:

t_a = 20 millisecs for a 10 dB gain reduction in A_1

t_{r1} = 3 secs for a 10 dB gain increase in A_1

t_{r2} = 1 min for a 2 dB gain increase in A_1

In tests using a variety of recorded telephone-speech excerpts at differing levels it was found that high-level signals were attenuated very quickly. Signals requiring a full 40 dB of attenuation were controlled in level within one second. When the level controller was required to increase the level of the signal by a full 40 dB the slewing time was about 10 to 15 secs depending on the nature of the signal.

A manually operated reset was provided on the equipment (see Fig. 13) so that the gain of the level controller could be set to maximum, hence reducing the slewing time for extremely low-level signals. In practice this facility was used infrequently to ensure that the first few sentences of a

new call were not lost when the incoming signal level was extremely low. For the level differences usually encountered, the device was quick to act and required no manual operation.

The only operational disadvantage of this automatic level controller was that the unwanted signal which appeared with the incoming telephone signals from the hybrid coil (see Fig. 10) was also controlled to the standard level. This was because the pause sensor did not discriminate between the incoming call and the unwanted signal. Hence the operation of an adaptive equaliser would be impaired. A possible remedy would be to derive additional control signals for the variable-gain amplifier A_1 from the studio signal which is applied to the telephone line; e.g. to inhibit the operation and switch to the long time constant when outgoing studio signals are present.

SMW/AH